

STORAGE DEVICE, DATA STORING METHOD AND CONTROL DEVICE

Publication number: JP2003241905

Publication date: 2003-08-29

Inventor: KANAI HIROKI

Applicant: HITACHI LTD

Classification:

- international: G06F12/08; G06F3/06; G06F12/00; G06F12/08;
 G06F3/06; G06F12/00; (IPC1-7): G06F3/06; G06F12/00;
 G06F12/08

- european:

Application number: JP20020039330 20020215

Priority number(s): JP20020039330 20020215

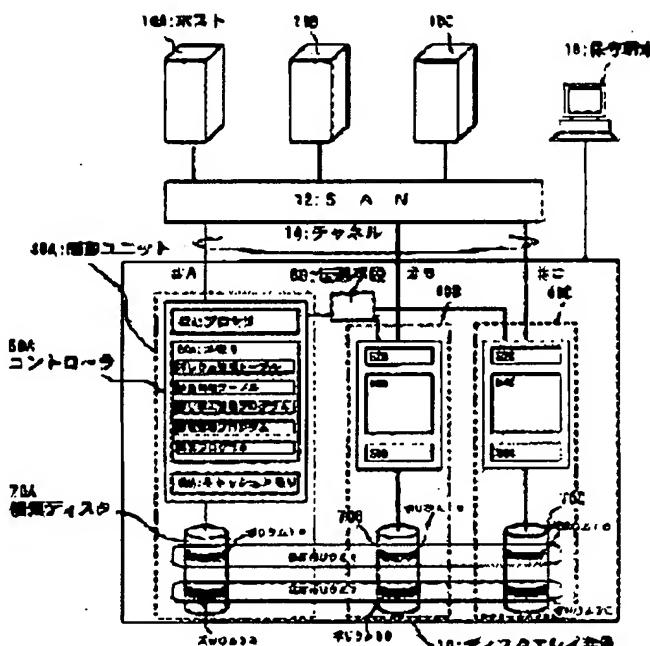
Report a data error here

Abstract of JP2003241905

PROBLEM TO BE SOLVED: To improve performance of a disk array device to solve a problem that on the so-called desk array device controlling storage devices and operating them as one storage device as a whole, when the performance of a transmission means transmitting data is low, data transfer between control units is slow and the entire performance of the disk array device is lowered.

SOLUTION: On this storage device provided with a data transmission means 112 connected with a plurality of units 140, a storing means 170 of a first unit stores copies of the data stored in the storing means of other units of units 140 and when the first unit of units 140 receives a read request from the outside for the data stored in the storing means of the other unit of units 140, the copies stored in the storing means of the first unit of the units 140 are transmitted to the outside. Accordingly, as the copies of the same data are stored in a plurality of the storing means, a frequency of access to the data transmission means 112 can be lessened.

COPYRIGHT: (C)2003,JPO



(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2003-241905

(P2003-241905A)

(43)公開日 平成15年8月29日(2003.8.29)

| (51) Int.Cl. ¹ | 識別記号 | F I | テ-マコト ² (参考) |
|---------------------------|-------|--------------|-------------------------|
| G 0 6 F 3/06 | 3 0 2 | G 0 6 F 3/06 | 3 0 2 Z 5 B 0 0 5 |
| | 3 0 1 | | 3 0 1 B 5 B 0 6 5 |
| 12/00 | 5 4 5 | 12/00 | 6 4 5 A 5 B 0 8 2 |
| 12/08 | 5 3 1 | 12/08 | 5 3 1 C |
| | 5 5 7 | | 5 5 7 |

審査請求 未請求 請求項の数 8 OL (全 22 頁)

(21)出願番号 特願2002-39330(P2002-39330)

(22)出願日 平成14年2月15日(2002.2.15)

(71)出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72)発明者 金井 宏樹

神奈川県小田原市中里322番地2号 株式
会社日立製作所 R&Dシステム事業部内

(74)代理人 100071283

弁理士 一色 錠輔 (外5名)

Fターム(参考) S8005 MM12

S8065 BA01 CA13 CA30 CC03 CC08

CE01 CH13

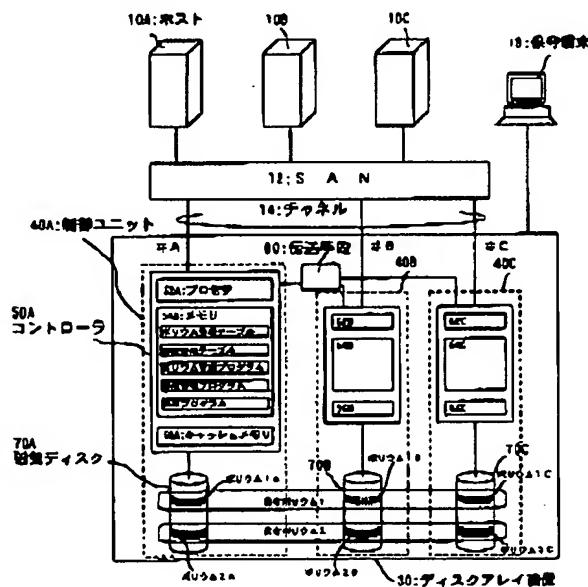
S8082 CA18 CA20 FA12

(54)【発明の名称】記憶装置、データ記憶方法及び制御装置

(57)【要約】 (修正有)

【課題】記憶装置をコントロールし全体として1つの記憶装置のように機能させるいわゆるディスクアレイ装置においてデータを送る伝送手段の性能が低いと、制御ユニット間のデータの転送が遅くディスクアレイ装置全体の性能が低下する。このようなディスクアレイ装置の性能を向上する。

【解決手段】複数の該ユニット140を接続するデータ伝送手段112を備える記憶装置であって、一のユニットの記憶手段170に、他のユニット140の記憶手段に記憶されたデータの複数を記憶し、該一のユニット140の制御装置が該他のユニット140の記憶手段に記憶されたデータの読み出し要求を外部から受信したとき、該一のユニット140の記憶手段に記憶した該複数を該外部に送信することを特徴とする。これにより、複数の記憶手段に同一のデータを複製して記憶させているので、データ伝送手段112へのアクセス頻度を少なくすることができる。



(2)

特開2003-241905

2

【特許請求の範囲】

【請求項1】 記憶手段と該記憶手段に対するデータの入出力を制御する制御装置とを有するユニットを複数備え、複数の該ユニットを接続するデータ伝送手段を備える記憶装置であって、

一のユニットの記憶手段に、他のユニットの記憶手段に記憶されたデータの複製を記憶し、

該一のユニットの制御装置が該他のユニットの記憶手段に記憶されたデータの読み出し要求を外部から受信したとき、該一のユニットの記憶手段に記憶した該複製を該外部に送信することを特徴とする記憶装置。

【請求項2】 請求項1に記載の記憶装置であって、前記一のユニットの記憶手段に記憶されたデータの格納位置と、前記他のユニットの記憶手段に記憶されたデータの格納位置とを対応付けた情報を有し、

該一のユニットの制御装置が該他のユニットの記憶手段に記憶されたデータの読み出し要求を前記外部から受信したとき、該情報に基づいて、該読み出し要求を受けたデータに対応する格納位置に記憶された前記複製のデータを該外部に送信することを特徴とする記憶装置。

【請求項3】 請求項1又は2に記載の記憶装置であつて、

前記他のユニットの制御手段は、該他のユニットの記憶手段に記憶された前記データに対する書き込み要求を前記外部から受信したとき、書き込み要求があった旨の通知を前記一のユニットに対して送信することを特徴とする記憶装置。

【請求項4】 請求項3に記載の記憶装置であって、前記他のユニットの制御手段が、前記外部に対して、書き込み処理の終了を通知するのは、

前記書き込み要求があった旨の通知を前記一の制御装置に送信した後であることを特徴とする制御装置。

【請求項5】 請求項1～4のいずれかに記載の記憶装置であつて、

前記他のユニットの記憶手段に記憶されたデータに対して書き込みを行うとき、

該他のユニットの制御装置は、前記一のユニットに対して、書き込みの内容を送信し、

該一のユニットの制御装置は、該他のユニットから、書き込みの内容を受信し、

該一のユニットの制御装置は、受信した書き込みの内容を前記複製に対して書き込むことを特徴とする記憶装置。

【請求項6】 記憶手段と該記憶手段に対するデータの入出力を制御する制御装置とを有するユニットを複数備え、複数の該ユニットを接続するデータ伝送手段を備える記憶装置のデータ記憶方法であって、

一のユニットの記憶手段に、他のユニットの記憶手段に記憶されたデータの複製を記憶し、

該一のユニットの制御装置が該他のユニットの記憶手段に記憶されたデータの読み出し要求を外部から受信したと

き、該一のユニットの制御装置が制御する記憶手段に記憶する該複製を該外部に送信することを特徴とするデータ記憶方法。

【請求項7】 記憶手段に対するデータの入出力を制御し、他の記憶手段のデータの入出力を制御する他の制御装置とデータ伝送手段を介して接続される制御装置であつて、

該他の記憶手段に記憶されたデータの複製を、該データ伝送手段を介して、自らが制御する該記憶手段に記憶し、

該他の記憶手段に記憶されたデータの読み出し要求を外部から受信したとき、該外部に対し、該複製を送信することを特徴とする制御装置。

【請求項8】 請求項7に記載の制御装置であつて、前記記憶手段を備えていることを特徴とする制御装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、複数の制御手段を制御する記憶装置の制御装置に関する。また、このような制御装置の制御方法及びこのような制御装置を備えた記憶装置に関する。

【0002】

【従来の技術】 近年、複数の記憶手段をコントロールし、全体として1つの記憶装置のように機能させる、いわゆるディスクアレイ装置が知られている。このようなディスクアレイ装置として、特開2001-256003号公報に記載されたような、複数のユニットを結合させてディスクアレイ装置を構成するものが提案されている。

【0003】 図15は、複数のユニットを結合させたディスクアレイ装置を備えたコンピュータシステムの概略図である。このシステムは、主として、ホスト110、SAN112及びディスクアレイ装置130を備えている。ディスクアレイ装置130は、複数の制御ユニット140とデータ伝送手段180とを備えている。各制御ユニット140は、コントローラ150と磁気ディスク170とを有している。コントローラ150内のメモリには、どのようなデータがどこに格納されているのかを対応させるためのテーブルがある。データ伝送手段180は、制御ユニット140同士を接続し、制御ユニット間での情報の伝達を可能としている。

【0004】 次に、ホスト110Aが制御ユニット140Aに対してデータの読み出しを要求する場合について説明する。ホスト110Aから読み出し要求を受信した制御ユニット140Aのコントローラ150Aのプロセサは、コントローラ150Aのメモリにアクセスし、テーブルに基づいて、要求されたデータがどこに格納されているかを調べる。要求されたデータが制御ユニット140Aのキャッシュメモリ又は磁気ディスク170Aに存在する場合は、そのデータをホスト110Aに送る。一

(3)

特開2003-241905

3

方、要求されたデータが他の制御ユニット（例えば、制御ユニット140Bのキャッシュメモリ又は磁気ディスク）に存在する場合は、データ伝送手段180を介して、そのデータを制御ユニット140Aのキャッシュメモリに転送し、そのデータをホスト110Aに送る。

【0005】

【発明が解決しようとする課題】しかし、要求されたデータが他の制御ユニットに存在する場合、要求されたデータは、このデータを格納する制御ユニットから、データ伝送手段180を介して、読み出し要求を受信した制御ユニットに送信され、その後に、ホスト110に転送されることになる。そのため、データ伝送手段180の性能が低いと、制御ユニット間のデータの転送が遅く、ホストへ要求されたデータを送るのが遅くなるので、ディスクアレイ装置全体としての性能が低くなる。一方、データ伝送手段180は高価であるため、制御ユニット間のデータの転送を早めるために高性能なデータ伝送手段を用いると、ディスクアレイ装置のコストが高くなる。

【0006】そこで、本発明は、このようなディスクアレイ装置等の記憶装置の性能を向上させることを目的とする。

【0007】

【課題を解決するための手段】上記の目的を達成するための主たる発明は、記憶手段と該記憶手段に対するデータの入出力を制御する制御装置とを有するユニットを複数備え、複数の該ユニットを接続するデータ伝送手段を備える記憶装置であって、一のユニットの記憶手段に、他のユニットの記憶手段に記憶されたデータの複製を記憶し、該一のユニットの制御装置が該他のユニットの記憶手段に記憶されたデータの読み出し要求を外部から受信したとき、該一のユニットの記憶手段に記憶した該複製を該外部に送信することを特徴とする記憶装置である。

【0008】

【発明の実施の形態】（実施形態1）

＜開示の概要＞本明細書及び添付図面の記載により、少なくとも、以下の事項が明らかとなる。記憶手段と該記憶手段に対するデータの入出力を制御する制御装置とを有するユニットを複数備え、複数の該ユニットを接続するデータ伝送手段を備える記憶装置であって、一のユニットの記憶手段に、他のユニットの記憶手段に記憶されたデータの複製を記憶し、該一のユニットの制御装置が該他のユニットの記憶手段に記憶されたデータの読み出し要求を外部から受信したとき、該一のユニットの記憶手段に記憶した該複製を該外部に送信することを特徴とする記憶装置。このような記憶装置によれば、複数の記憶手段に同一のデータを複製して記憶させているので、データ伝送手段へのアクセス頻度を少なくすることができる。その結果、データ読み出し時における記憶装置の性能を高めることができる。

【0009】また、かかる記憶装置において、前記一の

10

4

ユニットの記憶手段に記憶されたデータの格納位置と、前記他のユニットの記憶手段に記憶されたデータの格納位置とを対応付けた情報を有し、該一のユニットの制御装置が該他のユニットの記憶手段に記憶されたデータの読み出し要求を前記外部から受信したとき、該情報に基づいて、該読み出し要求を受けたデータに対応する格納位置に記憶された前記複製のデータを該外部に送信することが望ましい。このような記憶装置によれば、複数の記憶手段に同一のデータを複製して記憶させているので、データ伝送手段へのアクセス頻度を少なくすることができる。

20

【0010】また、かかる記憶装置において、前記他のユニットの制御手段は、該他のユニットの記憶手段に記憶された前記データに対する書き込み要求を前記外部から受信したとき、書き込み要求があった旨の通知を前記一のユニットに対して送信することが望ましい。このような記憶装置によれば、複製された同一のデータ間での整合性を維持することができる。

30

【0011】また、かかる記憶装置において、前記他のユニットの制御手段が、前記外部に対して、書き込み処理の終了を通知するのは、前記書き込み要求があった旨の通知を前記一の制御装置に送信した後であることが望ましい。このような記憶装置によれば、複製された同一のデータ間での整合性を維持することができる。

30

【0012】また、かかる記憶装置において、前記他のユニットの記憶手段に記憶されたデータに対して書き込みを行うとき、該他のユニットの制御装置は、前記一のユニットに対して、書き込みの内容を送信し、該一のユニットの制御装置は、該他のユニットから、書き込みの内容を受信し、該一のユニットの制御装置は、受信した書き込みの内容を前記複製に対して書き込むことが望ましい。このような記憶装置によれば、複製された同一のデータ間での整合性を維持することができる。

40

【0013】また、記憶手段と該記憶手段に対するデータの入出力を制御する制御装置とを有するユニットを複数備え、複数の該ユニットを接続するデータ伝送手段を備える記憶装置のデータ記憶方法であって、一のユニットの記憶手段に、他のユニットの記憶手段に記憶されたデータの複製を記憶し、該一のユニットの制御装置が該他のユニットの記憶手段に記憶されたデータの読み出し要求を外部から受信したとき、該一のユニットの制御装置が制御する記憶手段に記憶する該複製を該外部に送信することを特徴とするデータ記憶方法。このような制御方法によれば、複数の記憶手段に同一のデータを複製して記憶させているので、データ伝送手段へのアクセス頻度を少なくすることができる。その結果、データ読み出し時における記憶装置の性能を高めることができる。

50

【0014】また、記憶手段に対するデータの入出力を制御し、他の記憶手段のデータの入出力を制御する他の制御装置とデータ伝送手段を介して接続される制御装置

(4)

特開2003-241905

5

であって、該他の記憶手段に記憶されたデータの複製を、該データ伝送手段を介して、自らが制御する該記憶手段に記憶し、該他の記憶手段に記憶されたデータの読み出し要求を外部から受信したとき、該外部に対し、該複製を送信することを特徴とする制御装置。このような制御装置によれば、複数の記憶手段に同一のデータを複製して記憶させているので、データ伝送手段へのアクセス頻度を少なくすることができる。その結果、データ読み出し時における記憶装置の性能を高めることができる。

【0015】また、かかる制御装置において、前記記憶手段を備えていることが望ましい。このような制御装置によれば、記憶手段と一体化させてモジュールとすることができます。

【0016】<本システムの構成>図1は、本実施形態のコンピュータシステムの概略図である。前述の従来のコンピュータシステムと比較すると、特に、それぞれの制御ユニットの磁気ディスクに、同一のデータを多重して記憶させている点で異なっている。本システムは、主として、ホスト10、SAN12、保守端末16及びディスクアレイ装置30を備えている。

【0017】ホスト10は、上位装置として機能し、例えば汎用コンピュータやパーソナルコンピュータなどで構成される。また、ホスト10は、本システムでは、複数台設けられている。なお、ホスト10は、メインフレームとして複数の論理ブロックを有していても良いし、オープンシステムを構築するためのプラットフォームでも良い。また、複数のホスト10が、メインフレームとオープンシステムの混在系であって良い。

【0018】SAN12は、ストレージ・エリア・ネットワークである。SAN12は、ホスト10とディスクアレイ装置30との間に設けられ、ファイバチャネルなどのインターフェースを介して、両者の間のデータの転送を行う。また、SAN12は、ホスト10とディスクアレイ装置30との間に複数の転送経路（アクセスルート）を構築している。なお、一般に、SAN12はファイバチャネルを使用し、チャネル14も同様にファイバチャネルを使用する。ファイバチャネルは、チャネルのポート毎の識別番号としてワールドワイドネイム（WWN）を備えている。また、ファイバチャネルを用いる代わりに、イーサーネットを用いる場合もあるが、この場合も同様の識別子を備える。ファイバチャネルの上位プロトコルにスカジイ（SCSI）を用いた場合、ボリウム識別番号としてSCSI IDを備えている。例えばチャネルにファイバチャネル、その上位レイヤプロトコルにSCSIを用いる場合、ホスト10は、ポート番号としてWWNを用い、各チャネル上に配置したボリウムの識別番号としてSCSI IDを用いてアクセス先のボリウムを指定し、アクセス要求を行う。

【0019】保守端末16は、ディスクアレイ装置30の各種設定、管理又は監視を行うためのものである。そ

10

して、保守端末16には、このような機能を実行するためのプログラムが組み込まれている。保守端末16は、例えば、汎用コンピュータやパーソナルコンピュータなどで構成される。保守端末16は、LAN又は専用線により、ディスクアレイ装置30と接続している。なお、保守端末16には、ディスクアレイ装置30とLANなどを介して接続している端末に対し、リモート操作が可能な他の端末も含まれる。

【0020】ディスクアレイ装置30は、下位装置としての記憶装置として機能する。ディスクアレイ装置30は、ホスト10からデータの書き込み要求を受信したとき、ホスト10から受信した書き込みデータを記憶する。また、ディスクアレイ装置30は、ホスト10からデータの読み出し要求を受信したとき、記憶しているデータをホスト10に送信する。ディスクアレイ装置30は、複数の制御ユニット40とデータ伝送手段80とを備えている。

【0021】各制御ユニット40は、コントローラ50と磁気ディスク70とを有している。コントローラ50は、磁気ディスクに対するデータの入出力を制御する。各コントローラ50は、プロセサ52とメモリ54とキャッシュメモリ56とを有する。各メモリ54は、ボリウム管理テーブルと、排他管理テーブルと、ボリウム管理プログラムと、排他管理プログラムと、処理プログラムとを格納している（メモリ54B及びメモリ54Cも、不図示であるが、メモリ54Aと同様である）。これらのテーブル及びプログラムについては、後述する。磁気ディスク70は、記憶手段として機能し、種々の情報を記憶する。

20

【0022】本実施形態では、各制御ユニット40の各磁気ディスク70の一部のボリウムから共有ボリウムを構成し、共有ボリウムとして割り当たられた各磁気ディスクのボリウムに同一のデータを複製している。つまり、ある磁気ディスク70が、他の制御ユニットの磁気ディスクに格納されたデータと同一のデータを複製して記憶している。例えば、同図において、共有ボリウム1はボリウム1Aとボリウム1Bとボリウム1Cから構成され、ボリウム1Aとボリウム1Bとボリウム1Cは同一のデータを格納している。そして、ボリウム1Aに記憶されたデータはボリウム1Bとボリウム1Cとに複製を記憶し、ボリウム1Bに記憶されたデータはボリウム1Cとボリウム1Aとに複製を記憶し、ボリウム1Cに記憶されたデータはボリウム1Aとボリウム1Bとに複製を記憶する。なお、共有ボリウム2のボリウム2Aとボリウム2Bとボリウム2Cに関しても同様である。

30

【0023】データ伝送手段80は、制御ユニット40同士を接続し、制御ユニット間での情報の伝達を可能としている。データ伝送手段80は、例えばスイッチにより構成される。

40

【0024】なお、本実施形態では、コントローラ50

(5)

特開2003-241905

と磁気ディスク70が、1つのモジュールとして同一の筐体に構成されている。しかし、これに限られるものではない。例えば、コントローラ50と磁気ディスク70を別の筐体にしても良いし、他の形態であっても良い。要するに、コントローラ50と磁気ディスク70が、後述する機能を有していれば良い。

【0025】また、本実施形態では、1つの制御ユニット40は、1つのモジュールとしてディスクアレイ装置30の筐体の中に構成されている。しかし、これに限られるものではない。例えば、1つの制御ユニット40を1つの筐体として構成しても良いし、他の形態であっても良い。要するに、制御ユニット40が、下記に説明される機能を有していれば良い。

【0026】<ポリウム管理テーブル>図2は、図1のメモリ54内に格納されているポリウム管理テーブルの説明図である。なお、ポリウム管理テーブルは、各制御ユニット40のコントローラ50のメモリ54にそれぞれ格納されている。

【0027】ポリウム管理テーブルは、『共有ポリウム識別番号』、『チャネル番号』、『ポリウムの識別番号』及び『制御ユニットの識別番号』を関連付けたデータである。『共有ポリウム識別番号』とは、各磁気ディスクのポリウムに同一のデータを記憶させる際に、その共有単位毎につけられる識別子である。なお、共有ポリウム識別番号は、ディスクアレイ装置30内において、ユニークな識別子である。『チャネル番号』とは、ホストと接続するチャネル毎につけられる識別子である。例えば、チャネル番号として、ワールドワイドネイム(WWN)が用いられる。なお、チャネル番号は、ディスクアレイ装置30内において、ユニークな識別子である。

『ポリウムの識別番号』とは、ホストが認識するチャネル上のポリウム番号である。例えば、ポリウムの識別番号としてSCSI IDが用いられる。『制御ユニットの識別番号』とは、当該ポリウムを配置した制御ユニットの識別子である。

【0028】なお、ポリウム管理テーブルは、図に示された要素を必須の構成とするものではない。例えば、『共有ポリウム識別番号』と『ポリウム識別番号』を関連付けたデータであっても良いし、『アドレス番号』等の他の情報を関連付けたデータであっても良い。

【0029】<データの読み出し処理(排他処理なし)>次に、ホスト10が、データの読み出しを要求する場合について説明する。ただし、ここでは排他処理(後述)を行われていない状態とする。

【0030】データの読み出し要求が、ホスト10からチャネルを介して、制御ユニット40に送られる。読み出し要求を受信した制御ユニット40のコントローラ50のプロセサ52は、メモリ54にアクセスし、ポリウム管理テーブルを参照する。そして、ポリウム管理テーブルに基づいて、ホストが指定したポリウムが、共有ポリウ

50

8

ム内にあるかを調べる。要求されたデータが共有ポリウム内に存在する場合は、制御ユニット40の磁気ディスク70に同一のデータ(複製)が記憶されているので、そのデータをホスト10に送信する。なお、要求されたデータが、共有ポリウム内になく、且つ、他の制御ユニットに存在する場合は、データ伝送手段80を介して、そのデータを制御ユニット40Aのキャッシングメモリに転送し、そのデータをホスト10Aに送る。

【0031】具体例として、ホスト10Aが、チャネル#Aを介して、ポリウム1Bのデータの読み出しを要求する場合について説明する。

【0032】この場合、データの読み出し要求が、ホスト10Aからチャネル#Aを介して、制御ユニット40Aに送られる。コントローラ50Aのプロセサ52Aは、メモリ54Aに格納されているポリウム管理テーブルを参照する。そして、ポリウム管理テーブルに基づいて、ホスト10Aから指定されたポリウム1Bが共有ポリウム1として各制御ユニットの磁気ディスクに共有されていることが、認識される。さらに、共有ポリウム1のポリウム1Aに、ポリウム1Bと同一のデータが格納されていることも認識される。そこで、ディスクアレイ装置30は、ポリウム1Bのデータの読み出しを要求したホスト10Aに対して、ポリウム1Aに格納されているデータを送る。つまり、このような場合、データ伝送手段80を介さずに、ホストから要求されたデータを送ることができる。

【0033】なお、上記のようなデータの読み出し処理は、各制御ユニット40の各メモリ54内のポリウム管理プログラムや処理プログラムを実行することによって行われる。

【0034】本実施形態では、それぞれの制御ユニット40の磁気ディスク70に同一のデータを記憶させているので、データ伝送手段80へのアクセス頻度を少なくすることができる。その結果、データ読み出し時におけるディスクアレイ装置の性能を高めることができる。

【0035】なお、本実施形態では、ホスト10Aは制御ユニット40Aにデータを要求するだけで制御ユニット40Bのデータが送られてくるので、ディスクアレイ装置が複数の制御ユニットを備えていても、全体として1つの制御ユニットと同様にみなすことができる。したがって、ディスクアレイ装置の制御ユニットの増設が容易になるので、小規模な構成から大規模な構成までのスケーラビリティのある構成のディスクアレイ装置を提供することができる。

【0036】<排他管理テーブル>上記のように、データの読み出しは、各磁気ディスクが同一のデータを格納しているので、独立して行うことができる場合がある。しかし、データの書き込みの場合、共有ポリウム内のデータの一貫性を保持する必要がある。そこで、本実施形態では、共有ポリウムに格納されたデータの整合性を保つた

(6)

9

め、排他的な処理を行うための排他管理テーブルを設けている。

【0037】図3は、図1のメモリ54内に格納されている排他管理テーブルの説明図である。なお、排他管理テーブルは、各制御ユニット40のコントローラ50のメモリ54にそれぞれ格納されている。

【0038】排他管理テーブルは、『共有ポリウム識別番号』、『排他対象アドレス』、『更新データ格納ユニット』、『キャッシュアドレス』及び『データ更新有効フラグ』を関連付けたデータである。『共有ポリウム識別番号』は、前述のポリウム管理テーブルと同様である。『排他対象アドレス』とは、排他処理の対象として指定するアドレス先である。排他対象アドレスとして排他管理テーブルに登録されると、後述するように、そのアドレスへのデータの書き込みが制限される。『更新データ格納ユニット』とは、ホストから書き込み要求があったときに、その書き込みデータ（更新データ）を記憶したキャッシュメモリを有する制御ユニットの識別子である。

【0039】例えば、ホスト10から書き込み要求があり、当該書き込みデータをキャッシュメモリ56Aに記憶したときは、『更新データ格納ユニット』には制御ユニット40Aを識別するためのデータが登録される。『キャッシュアドレス』は、キャッシュメモリ56に書き込みデータを記憶したときのキャッシュメモリ56上のアドレス先である。『更新フラグ』は、データの更新が可能かどうかを識別するためのフラグである。更新フラグが有効であるときは、排他管理テーブルの例えば『更新データ格納ユニット』や『キャッシュアドレス』等のデータを変更することが可能である。また、更新フラグが無効であるときは、排他管理テーブルのデータの変更は禁止される。なお、後述する通り、更新フラグが無効であるときは、各制御ユニットの磁気ディスクに同一データを更新処理している最中である。

【0040】なお、排他管理テーブルは、共有ポリウム内の同一データの整合性を保持するため、後述する通り、各制御ユニットにおいて同期をとっている。

【0041】<データの書き込み処理>次に、ホスト10が、データの書き込みを要求する場合について説明する。図4は、データの書き込み処理に関するフロー図である。

【0042】データの書き込み要求が、ホスト10からチャネルを介して、制御ユニット40に送られる（S101）。書き込み要求を受信した制御ユニット40のコントローラ50のプロセサ52は、メモリ54にアクセスし、排他管理テーブルを参照する（S102）。そして、排他管理テーブルに基づいて、ホストが書き込み先として指定したアドレスが、排他中であるかを調べる（S103）。書き込み先が排他管理テーブルに登録されていれば、当該書き込み先が排他中であると認識されるので、排他管理テーブルの更新処理（後述）が行われる（S200）。書き込み先が排他管理テーブルに登録されていな

特開2003-241905

10

ければ、当該書き込み先は排他中ではないと認識されるので、他の制御ユニットに書き込み要求があった旨を通知し、全ての制御ユニット40の排他管理テーブルに当該書き込み先を登録する（S104）。この結果、例えば、他のホストから同じアドレスを指定する書き込み要求があつても、上記S103において、排他中であると認識されることになる。また、ホストに書き込み処理の終了を報告する前に、他の制御ユニットに書き込み要求があつた旨を通知し、全ての制御ユニット40の排他管理テーブルの登録を行っているので、各制御ユニット40の間で排他管理テーブルの同期をとることができる。次に、書き込み要求を受信した制御ユニット40のキャッシュメモリ56に書き込みの要求があつたデータ（書き込みデータ）を記憶する（S105）。キャッシュメモリ56に書き込みデータを記憶したら、ホストに対して、書き込み処理の終了を報告する（S106）。その後、磁気ディスクの指定アドレスにキャッシュメモリ56に記憶された書き込みデータを書き込むとともに、他の制御ユニットの磁気ディスクにも同一のデータを書き込む（共有ポリウム更新処理（S107））。これにより、共有ポリウム内のデータの一貫性が保持される。なお、各制御ユニットの磁気ディスクに同一データを書き込む処理については、後述する。

【0043】具体例として、ホスト10Aが、ポリウム1Aの指定アドレスに対して、データの書き込みを要求する場合について説明する。

【0044】この場合、データの書き込み要求が、制御ユニット40Aに送られる。プロセサ52Aは、メモリ54Aに格納されている排他管理テーブルを参照する。そして、排他管理テーブルに基づいて、ポリウム1Aの指定アドレスに該当する共有ポリウムのアドレスが、排他中であるかを調べる。なお、排他中であるときの処理は、後述する。当該アドレスが排他中でなければ、書き込み要求があつた旨を他の制御ユニット40B、40Cに通知する。そして、全ての制御ユニット40A～40Cの排他管理テーブルに対し、『共有ポリウム識別番号』としてポリウム1Aが所属する共有ポリウム1を登録し、『排他対象アドレス』としてポリウム1Aの指定アドレスに該当する共有ポリウムのアドレス（例えば0x0000000016-0x0000000019）を登録し、『更新データ格納ユニット』としてポリウム1Aが所属する制御ユニット40Aを登録し、『キャッシュアドレス』として制御ユニット40Aのキャッシュメモリ56Aのアドレス（書き込みデータを一時記憶するアドレス）を登録し、『更新フラグ』を無効として登録する。全ての制御ユニット40A～40Cの排他管理テーブルに登録を行っているので、各制御ユニットの排他管理テーブルの同期をとができる。次に、キャッシュメモリ56Aの所定のアドレス（0x12340016-0x12340019）に書き込みデータを記憶する。キャッシュメモリ56Aに書き込みデータを記憶した

50

(7)

特開2003-241905

11

ら、ホストに対して、書き込み処理の終了を報告する。その後、ポリウム1Aの指定アドレスに対して、キャッシングメモリ56Aに記憶された書き込みデータを書き込む。さらに、共有ポリウム1内のデータの一貫性を保持するため、後述する通り、同一の書き込みデータが、ポリウム1Bとポリウム1Cの所定のアドレスに書き込まれる(つまり、ポリウム1Aのデータの複製が、ポリウム1Bとポリウム1Cに記憶される)。

【0045】なお、上記のようなデータの書き込み処理は、各制御ユニット40の各メモリ54内の排他管理プログラムや処理プログラムを実行することによって行われる。

【0046】本実施形態では、排他管理テーブルを設けることにより、共有ポリウムに格納されたデータの整合性を保つことができる。また、キャッシングメモリに書き込みデータを記憶した後にディスクアレイ装置がホストに対して書き込み処理の終了を報告するため、データ伝送手段を介して書き込みデータを送信する前にディスクアレイ装置がホストに対して報告を行うので、ホストに対する応答が速くなる。

【0047】<排他管理テーブルの更新処理>次に、排他管理テーブルの更新処理について説明する。排他管理テーブルの更新処理は、上述のS103において書き込み先として指定されたアドレスが排他中であった場合に行われる。図5は、排他管理テーブルの更新処理に関するフロー図である。

【0048】書き込み先が排他管理テーブルに登録されているので(S103)、排他管理テーブルの更新処理に移行する(S200)。制御ユニット40のプロセサ52は、排他管理テーブルの登録内容のうち、更新フラグを調べる(S201)。更新フラグが有効であれば、排他管理テーブルの『更新データ格納ユニット』として登録されている制御ユニット40にアクセスし、更新要求を行う(S202)。なお、制御ユニット40へのアクセスは、プロセサ52間の通信等で行う。そして、『更新データ格納ユニット』として登録されている制御ユニット40から、排他管理テーブルの更新が可能であるとの応答があれば(S203)、全ての制御ユニット40の排他管理テーブルに、新たに書き込み先に対応する『更新データ格納ユニット』と『キャッシングアドレス』とを登録する(S204)。一方、S201において更新フラグが無効である場合、又は、S203において排他管理テーブルの更新が不可である場合、排他管理テーブルの登録内容を変更することができないので、共有ポリウムの更新処理(後述)が終了するまで待機する(S205)。共有ポリウム更新処理が終了すれば、後述する通り、更新処理が終わった書き込み要求に関するデータが削除されているので、排他されていた書き込み要求を新たに排他管理テーブルに登録することができる(S206)。なお、S204の排他管理テーブルの変更や、S

10

12

206の排他管理テーブルの登録は、各制御ユニット40の排他管理テーブルの同期をとるために、全ての制御ユニット40の排他管理テーブルに行う。

【0049】具体例として、排他管理テーブルが図3に示す通りの状態であるときに、ホスト10Aが、ポリウム2Aの指定アドレスに対して、データの書き込みを要求する場合について説明する。ただし、ポリウム2Aの指定アドレスに対応する共有ポリウムのアドレスは、排他対象アドレスとなっているものとする。

【0050】この場合、制御ユニット40Aのプロセサ52Aは、メモリ54Aの排他管理テーブルを参照する。書き込み先が排他管理テーブルに登録されているので、排他管理テーブルの更新処理に移行する。プロセサ52Aは、排他管理テーブルの登録内容のうち、更新フラグを調べ、更新フラグが有効であることを確認する。そこで、制御ユニット40Aのプロセサ52Aは、『更新データ格納ユニット』に登録されている制御ユニット40Cのプロセサ52Cに対し、更新要求を行う。プロセサ52Aがプロセサ52Cから排他管理テーブルの更新が可能との応答を受けたときは、全ての制御ユニット40A~40Cの排他管理テーブルに対し、『更新データ格納ユニット』の『40C』を『40A』に登録内容を変更し、『キャッシングアドレス』の登録内容も制御ユニット40Aのキャッシングメモリ56Aに対応するように変更する。

【0051】なお、上記のような排他管理テーブルの更新処理は、各制御ユニット40の各メモリ54内の排他管理プログラムや処理プログラムを実行することによって行われる。

【0052】<共有ポリウムの更新処理>次に、上述のS107の共有ポリウムの更新処理について説明する。共有ポリウムの更新処理とは、上述のS105のキャッシングメモリ56に記憶された書き込みデータを、各制御ユニットのそれぞれの磁気ディスクに同一データを書き込む処理である。つまり、共有ポリウムの更新処理は、共有ポリウム内のデータの一貫性を保持するための処理である。図6は、共有ポリウムの更新処理に関するフロー図である。

【0053】共有ポリウムの更新処理が開始されると、40全ての制御ユニットの排他管理テーブルの更新フラグが無効となる(S301)。これにより、これから共有ポリウムの更新処理を行うデータは、更新処理を行っている間に、新たに更新されることはない。次に、排他管理テーブルの『更新データ格納ユニット』に登録されている制御ユニット40が、他の制御ユニットにデータ伝送手段80を介して、キャッシングメモリ56に記憶されている書き込みデータを送信する(S302)。これにより、全ての制御ユニット40のキャッシングメモリ56には、共有ポリウムに書き込みを行うための同一のデータが記憶される(S303)。そして、各制御ユニット40

(8)

13

は、それぞれの磁気ディスク70に書き込みデータの記憶を行う。この結果、各制御ユニットの磁気ディスク70に同一のデータが記憶されることになる。その後、全ての制御ユニットの排他管理テーブルから、この書き込み処理に関するデータを削除する(S304)。これにより、この書き込み要求の対象となった共有ポリウムのアドレスは、排他処理の対象から外れることになる。

【0054】本実施形態では、排他管理テーブルに更新フラグを設けたので、各制御ユニットの磁気ディスクに同一データを更新処理している最中に、当該データに対して更新処理が発生した場合でも、データの整合性を保つことができる。

【0055】<データの読み出し処理(排他処理あり)>次に、ホスト10が、データの読み出しを要求する場合について説明する。図7は、データの読み出し処理に関するフロー図である。

【0056】データの読み出し要求が、ホスト10からチャネルを介して、制御ユニット40に送られる(S401)。読み出し要求を受信した制御ユニット40のコントローラ50のプロセサ52は、メモリ54にアクセスし、排他管理テーブルを参照する(S402)。そして、排他管理テーブルに基づいて、ホストが読み出し先として指定したアドレスが、排他中であるかを調べる(S403)。書き込み先が排他管理テーブルに登録されていなければ、当該読み出し先は排他中ではないと認識される。この場合の読み出し処理は、排他処理の無いときのデータの読み出し処理として、既に説明済みなので省略する(S404)。読み出し先が排他管理テーブルに登録されていれば、当該読み出し先が排他中であると認識される。この場合、排他管理テーブルの『更新データ格納ユニット』として登録された制御ユニットのキャッシュメモリが、更新する書き込みデータ(すなわち最新のデータ)を格納している。したがって、排他管理テーブルの『更新データ格納ユニット』に登録された制御ユニットのキャッシュメモリから、データ伝送手段80を介して、読み出し要求を受信した制御ユニットのキャッシュメモリに読み出しだデータをコピーする(S411)。なお、読み出しだデータをコピーする主体は、読み出し要求を受信した制御ユニットでも良いし、排他管理テーブルの『更新データ格納ユニット』に登録された制御ユニットでも良い。キャッシュメモリに読み出しだデータをコピーしたら、ホストに対して、読み出しだデータを送信し(S405)、読み出し処理を終了する。

【0057】具体例として、ホスト10Aが、ポリウム1Aの指定アドレスに対して、データの読み出しを要求する場合について説明する。ただし、ポリウム1Aの指定アドレスに対応する共有ポリウムのアドレスは、排他対象アドレスとなっているものとする。

【0058】この場合、データの読み出し要求が、制御ユニット40Aに対して送られる。プロセサ52Aは、メ

特開2003-241905

14

モリ54Aに格納されている排他管理テーブルを参照する。そして、排他管理テーブルに基づいて、ポリウム1Aの指定アドレスに該当する共有ポリウムのアドレスが、排他中であるか調べる。ここでは図3に示す通りに排他中であるものとする。そこで、プロセサ52Aは、排他管理テーブルの『更新データ格納ユニット』に登録されている制御ユニット40Bのプロセサ52Bに対し、キャッシュメモリ56Bの所定のアドレス(0x5678000-0x56781024)に格納されているデータのコピーを要求する(なお、データのコピーは、プロセサ52Aがプロセサ52Bに要求することに限られるものではなく、プロセサ52Aが直接キャッシュメモリ56Bから読み出しを行うことによって実行しても良い。要は、キャッシュメモリ56Bの所定のデータが、キャッシュメモリ56Aにコピーできれば良いのである)。キャッシュメモリ56Bからキャッシュメモリ56Aに当該データがコピーされたら、プロセサ52Aは、ホスト10Aに対して、読み出しだデータを送信し、読み出し処理を終了する。

【0059】なお、上記のような読み出し処理は、各制御ユニット40の各メモリ54内のポリウム管理テーブル、排他管理プログラム及び処理プログラムを実行することによって行われる。

【0060】本実施形態によれば、排他中の読み出し処理であっても、最新のデータを読み込むことが可能であり、データの整合性を保つことができる。

【0061】<共有ポリウムの作成>次に、保守端末16が、ディスクアレイ装置30に対して、共有ポリウムの設定を行う場合について説明する。図8は、共有ポリウムの設定に関するフロー図である。

【0062】保守端末16が共有ポリウムの作成を要求すると、保守端末16で入力されたデータに基づいて、指定された制御ユニットのポリウム管理テーブルに、共有ポリウム識別番号を登録する(S501)。そして、制御ユニット40に対して、共有ポリウムを構成するポリウムを割り当てる(S502)。その後、作成したポリウムに関する情報を、ポリウム管理テーブルに登録する(S503)。次に、作成した共有ポリウムが、新規であるかどうかを判断する(S504)。作成した共有ポリウムが新規であれば、さらに他の制御ユニット40に対して共有ポリウムとしてポリウムを割り当てるかを判断する(S506)。作成した共有ポリウムが既に存在するのであれば、同じ共有ポリウムとなる他のポリウムから、この作成処理によって新たに割り当られたポリウムに同じデータをコピーする(複製を記憶する)(S505)。コピー終了後、さらに他の制御ユニット40に対して共有ポリウムとしてポリウムを割り当てるかを判断する(S506)。ポリウムの割り当てが終了すれば、共有ポリウム作成処理は終了する。

【0063】図9は、共有ポリウムの設定を保守端末1

(9)

15

6で行う場合の作成画面である。図9(A)は作成時の画面900Aを示し、図9(B)は作成結果の画面900Bを示している。

【0064】画面900Aにおいて、901は、共有ボリュームの一部となるボリュームを設定する領域である。ここでは、チャネル番号『#B』に対し、『100GB』のボリュームの設定を行っている。また、画面900Aにおいて、902は、共有ボリュームを設定又は特定する領域である。ここでは、『共有ボリューム2』を設定している。ここで入力された共有ボリュームに基づいて、上述のS504の判断を行うことになる。また、画面900Aにおいて、903は、設定内容の確定を行う領域である。領域903において、『NO』を選択すれば、領域901及び領域902で設定した内容がキャンセルされる。また、領域903において、『YES』を選択すれば、領域901及び領域902で設定した内容で、共有ボリュームの作成処理が行われる。

【0065】画面900Bは、領域903において『YES』を選択した後に表示される結果表示画面である。画面900Bの中央上段は、画面900Aにおいて作成した新規のボリュームに関する情報を表示している。ここでは、『チャネル#B』に対応して『ボリューム2B』が新たに作成されたことを示している。また、画面900Bの中央下段は、共有ボリュームに関する情報を表示している。ここでは、既に『共有ボリューム2』が存在しており、この共有ボリュームに対して『ボリューム2A』と『ボリューム2C』が既に割り当てられていたので、共有ボリューム2と対象となるボリューム(ボリューム2A~2C)を関連付けて表示している。ただし、結果表示画面は、これに限られるものではない。例えば、図2に示すようなボリューム管理テーブルをそのまま表示するような形態であっても良い。

【0066】なお、共有ボリュームが既に設定されている場合であって、画面900Aの領域901において新たに作成するボリュームの容量が、この共有ボリュームに対応する容量と異なる場合は、エラーが表示される。

【0067】上記の共有ボリュームの作成処理は、端末装置16によって行われるものとして説明されたが、これに限られるものではない。例えば、上記の共有ボリュームの作成処理をホスト10等から行っても良い。また、上記のような共有ボリュームの作成処理は、各制御ユニット40の各メモリ54内のボリューム管理テーブルや処理プログラムを実行することによって行われる。

【0068】(他の実施形態)以上、一実施形態に基づき、本発明に係る記憶装置等を説明した。しかし、上述の実施形態は、本発明の理解を容易にするためのものであって、本発明を限定して解釈するためのものではない。すなわち、本発明は、その趣旨を逸脱することなく、変更、改良され得ると共に、本発明にはその等価物が含まれることは言うまでもない。特に、以下に述べる

特開2003-241905

16

実施形態であっても、本発明に係る記憶装置等に含まれるものである。

【0069】<全体構成について>図10は、別の実施形態のコンピュータシステムの概略図である。このように、ホストとディスクアレイ装置30との間を、スイッチを主体とした中継器によって構成しても良い。また、これらに限られるものではなく、例えば、スイッチをループ構成にても良く、他の構成でも良い。また、同図に示す通り、各制御ユニットのボリュームが、複数の磁気ディスクにまたがって構成されても良い。さらに、各制御ユニットのボリュームが、RAIDを構成するようにしても良い。

【0070】また、図11は、さらに別の実施形態のコンピュータシステムの概略図である。このように、スイッチを介さずに、ホストとディスクアレイ装置との間を接続する構成であっても良い。従来のディスクアレイ装置は、特に、このような接続を行うと、データ伝送手段80のアクセス頻度が多くなるので、ディスクアレイ装置の性能が低下していた。しかし、本実施形態のディスクアレイ装置であれば、データ伝送手段80のアクセス頻度を少なくすることができるので、このようなホストとの接続を行っても、ディスクアレイ装置の性能の低下を抑えることができる。

【0071】<コントローラについて>前述の実施形態では、コントローラ50の概略を説明したが、これに限られるものではない。例えば、コントローラ50は、以下のよう構成であっても良い。要は、制御ユニット40が、前述の実施形態において説明した機能を果たせば良いのである。

【0072】図12は、別の実施形態のコントローラ50の構成の概略図である。前述の実施形態と同じ構成要素については、本実施形態では説明を省略する。同図において、コントローラ50は、チャネル制御部62と、制御メモリ54と、キャッシュメモリ56と、ディスク制御部64とを有する。チャネル制御部62は、ホスト10に対する入出力を制御する。ディスク制御部64は、ディスクに対する入出力を制御する。本実施形態のように、コントローラ50が、複数のチャネルを制御し、複数の磁気ディスクを制御するような形態であっても良い。

【0073】図13は、さらに別の実施形態のコントローラの構成の概略図である。図12と比較すると、キャッシュメモリが『ボリューム管理テーブル』と『排他管理テーブル』とを記憶している点で異なる。このような構成であっても、コントローラは、前述の実施形態と同様の機能を果たすことができる。

【0074】<データ伝送手段について>また、図14は、別の実施形態のデータ伝送手段の構成図である。前述の実施形態と比較すると、『ボリューム管理テーブル』と『排他管理テーブル』が、データ伝送手段に格納され

50

(10)

特開2003-241905

17

ている点が異なる。このように、データ伝送手段にボリューム管理テーブルや排他管理テーブルを設ければ、テーブルの一元管理が可能となるので、複数のテーブルの同期をとらずにすみ、処理が簡略化できる。

【0075】<テーブルについて>前述の実施形態において、ボリューム管理テーブル(図2)及び排他管理テーブル(図3)について説明したが、これらのテーブルは、この構成に限られるものではない。例えば、他の項目を備えても良いし、適宜項目を削除しても良い。要するに、コントローラが、上述した実施形態の機能を果たすことができれば良いのである。

【0076】<排他管理テーブルの更新について>前述の実施形態では、排他管理テーブルの更新処理において、『更新データ格納ユニット』と『キャッシュアドレス』とを登録していた(S204参照)。しかし、これに限られるものではない。要するに、共有ボリューム内に格納されている同一データの整合性が取れるのであれば、良いのである。例えば、排他中に新たな書き込みがあった場合、既に登録されているキャッシュアドレスに格納されているデータを新たなデータに書き替える、排他管理テーブルの『更新データユニット』や『キャッシュアドレス』の登録内容はそのままであっても良い。

【0077】

【発明の効果】本発明によれば、複数の記憶手段に同一のデータを複製して記憶させているので、データ伝送手段へのアクセス頻度を少なくすることができます。その結果、データ読み出し時における記憶装置の性能を高めることができる。

【図面の簡単な説明】

【図1】本実施形態のコンピュータシステムの概略図である。

【図2】図1のメモリ54内に格納されているボリューム管理テーブルの説明図である。

【図3】図1のメモリ54内に格納されている排他管理テーブルの説明図である。

【図4】データの書き込み処理に関するフロー図である。*

【図2】

| 共有ボリューム 識別番号 | チャネル番号 | ボリュームの 識別番号 | 記憶ユニット 識別番号 |
|-----------------|--------|----------------|----------------|
| 共有ボリューム 1 | #A | ボリューム1A | 40A |
| | #B | ボリューム1B | 40B |
| | #C | ボリューム1C | 40C |
| 共有ボリューム 2 | #A | ボリューム2A | 40A |
| | #B | ボリューム2B | 40B |
| | #C | ボリューム2C | 40C |

18

* 【図5】排他管理テーブルの更新処理に関するフロー図である。

【図6】共有ボリュームの更新処理に関するフロー図である。

【図7】データの読み出し処理に関するフロー図である。

【図8】共有ボリュームの設定に関するフロー図である。

【図9】共有ボリュームの設定を保守端末16で行う場合の作成画面である。

10 【図10】別の実施形態のコンピュータシステムの概略図である。

【図11】さらに別の実施形態のコンピュータシステムの概略図である。

【図12】別の実施形態のコントローラ50の構成の概略図である。

【図13】さらに別の実施形態のコントローラの構成の概略図である。

【図14】別の実施形態のデータ伝送手段の構成図である。

20 【図15】従来のディスクアレイ装置を備えたコンピュータシステムの概略図である。

【符号の説明】

- 10 ホスト
- 12 SAN
- 14 チャネル
- 16 保守端末
- 30 ディスクアレイ装置
- 40 制御ユニット
- 50 コントローラ
- 52 プロセッサ
- 54 メモリ
- 56 キャッシュメモリ
- 62 チャネル制御部
- 64 ディスク制御部
- 70 磁気ディスク
- 80 データ伝送手段

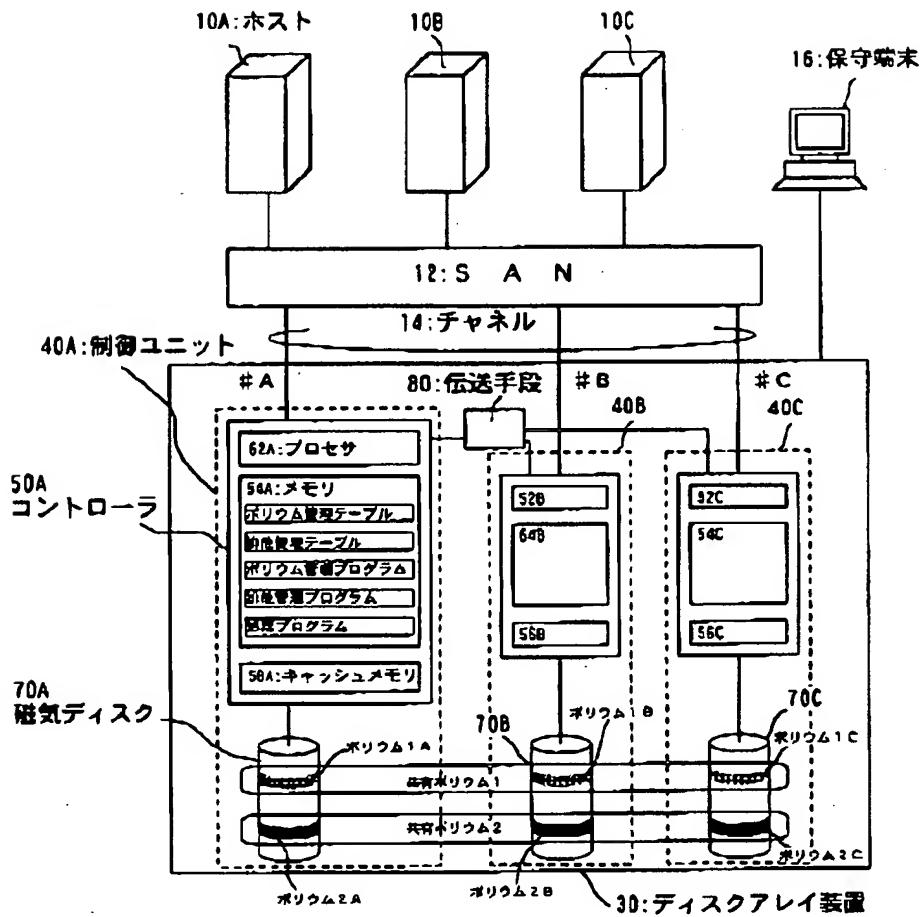
【図3】

| 共有ボリューム 識別番号 | 接続対象 アドレス | 更新データ 格納ユニット アドレス | キャッシュ アドレス | 更新フラグ |
|---------------------------|--------------|-------------------------|---------------|-------|
| 共有ボリューム0x0000000016- 1 | 0x0000000016 | 40A | 0x123400016- | 有効 |
| 共有ボリューム0x24850000- 2 | 0x2485000020 | 40C | 0x12370000- | 有効 |
| 共有ボリューム0x00200000- 1 | 0x0020000024 | 40B | 0x56780000- | 無効 |

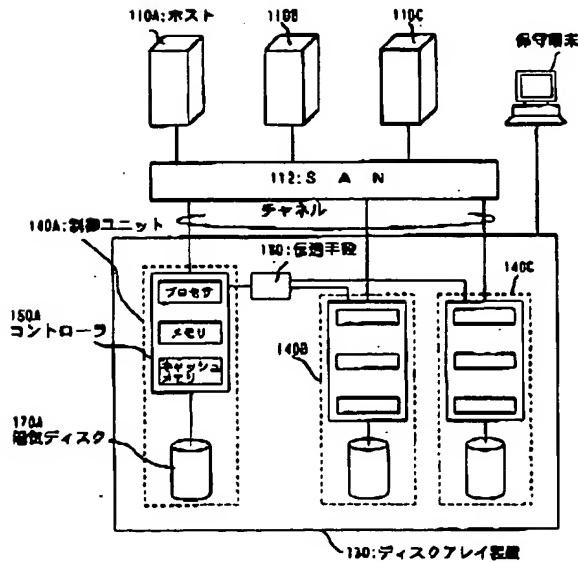
(11)

特開2003-241905

[1]



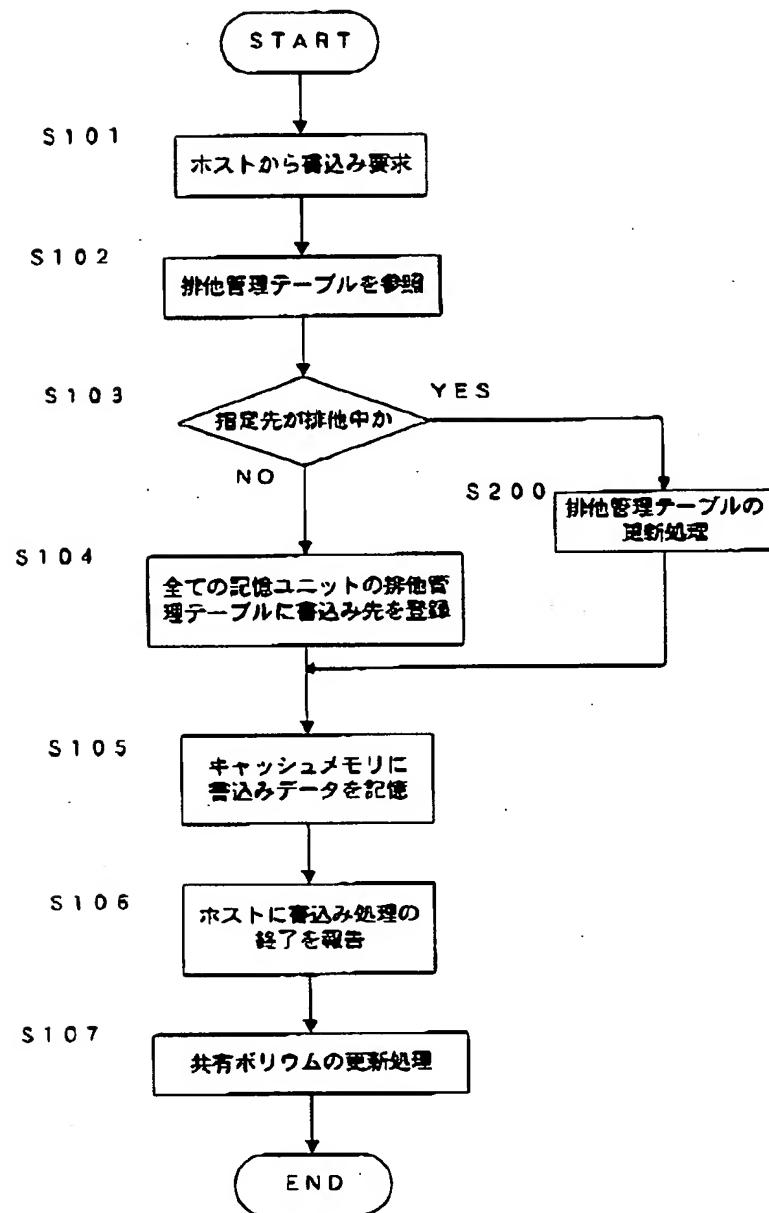
[1 5]



(12)

特開2003-241905

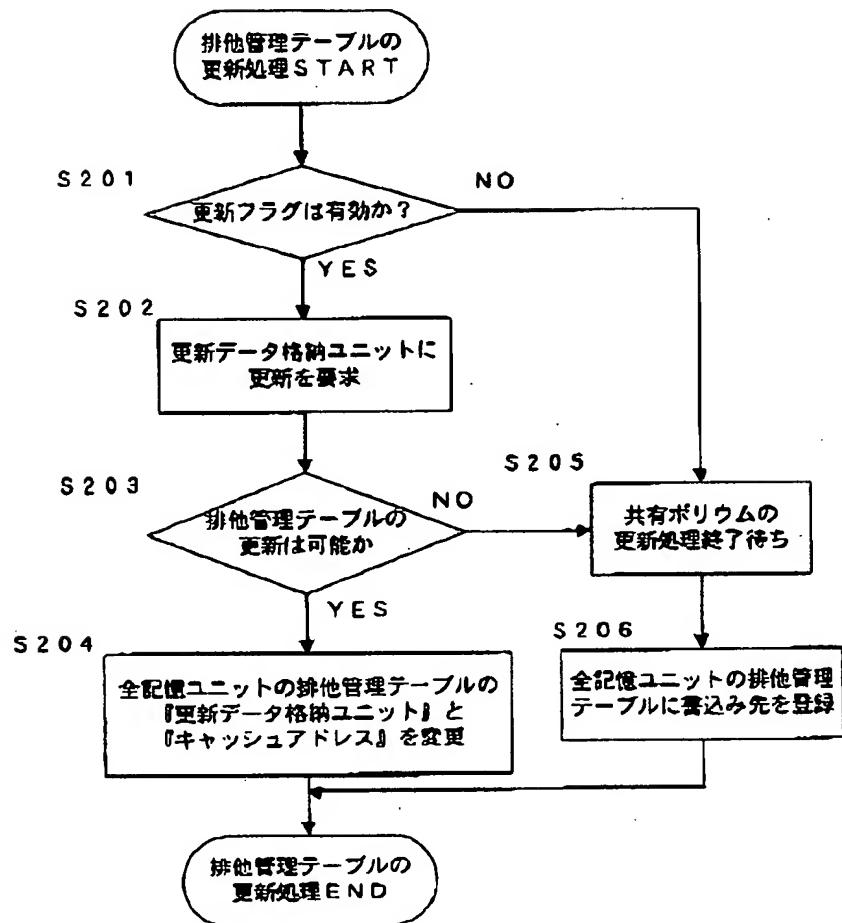
【図4】



(13)

特開2003-241905

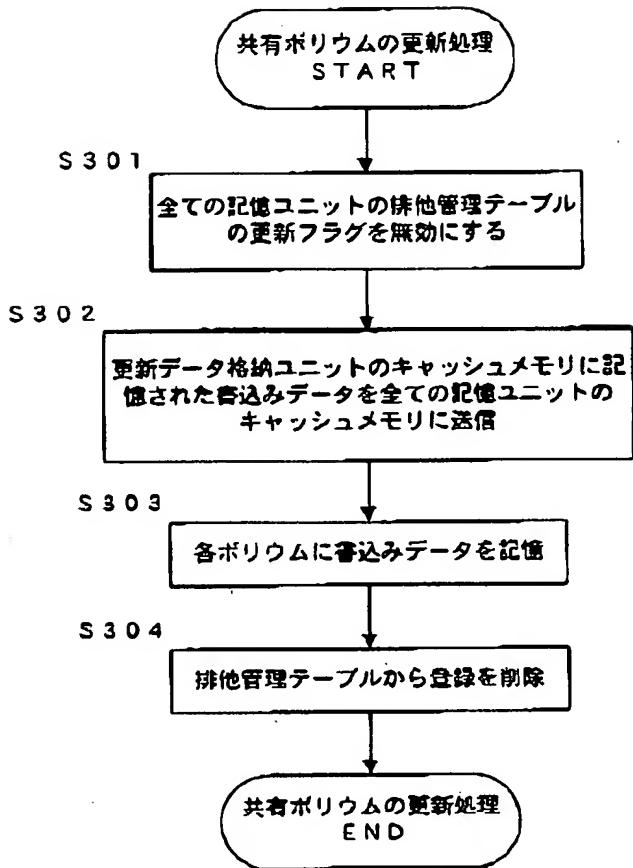
【図5】



(14)

特開2003-241905

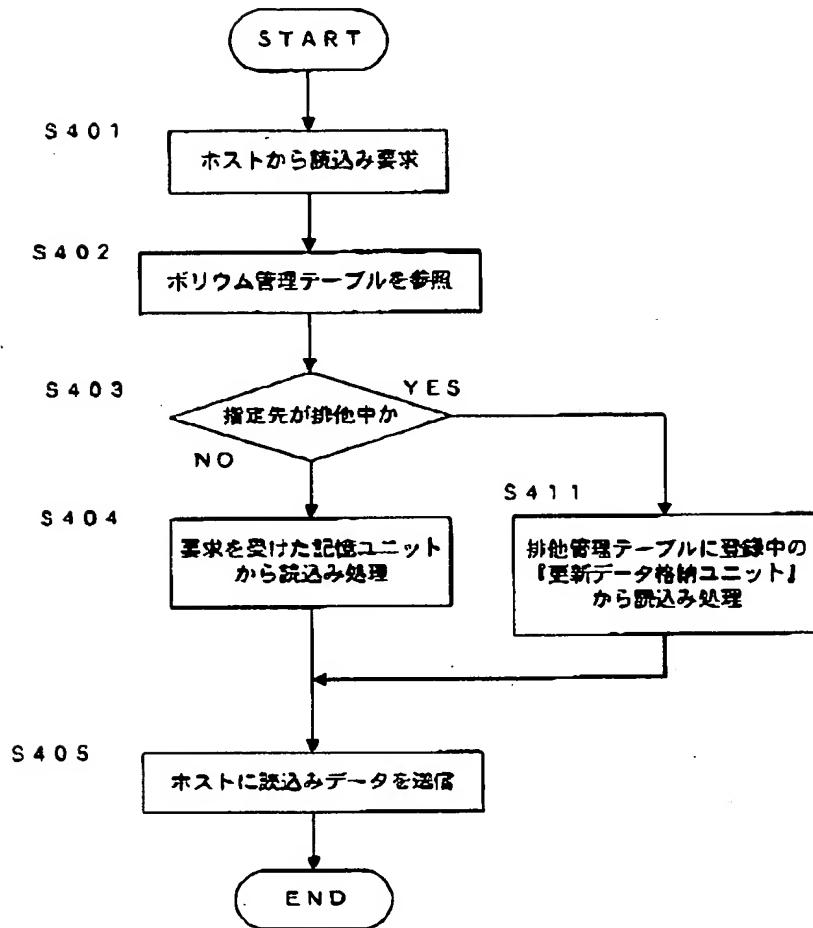
【図6】



(15)

特開2003-241905

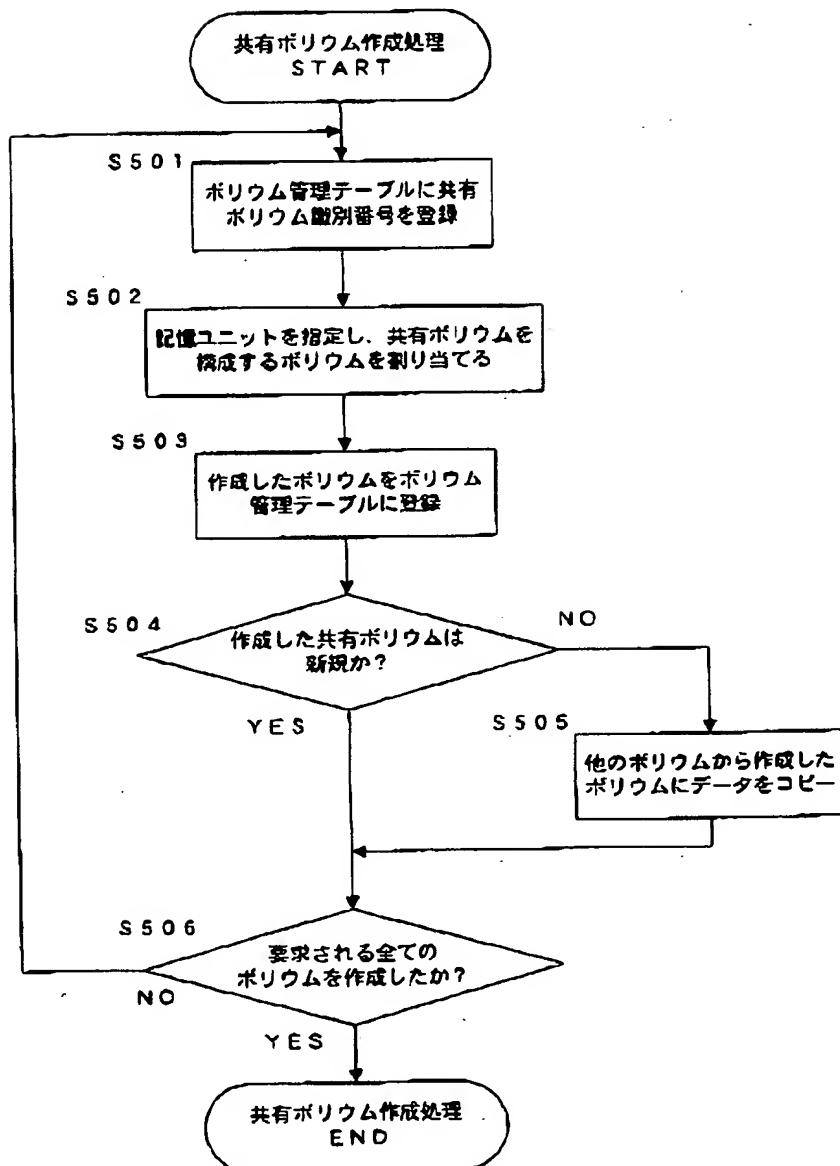
[図7]



(16)

特開2003-241905

[図8]



(17)

特開2003-241905

【図9】

(A) 900A

作成するボリュームの情報を登録してください。

901 チャネル番号とボリューム容量の設定 #B 100GB

902 共有ボリュームの設定 共有ボリューム 2

903 ボリューム作成を要求しますか？ YES NO

(B) 900B

登録結果

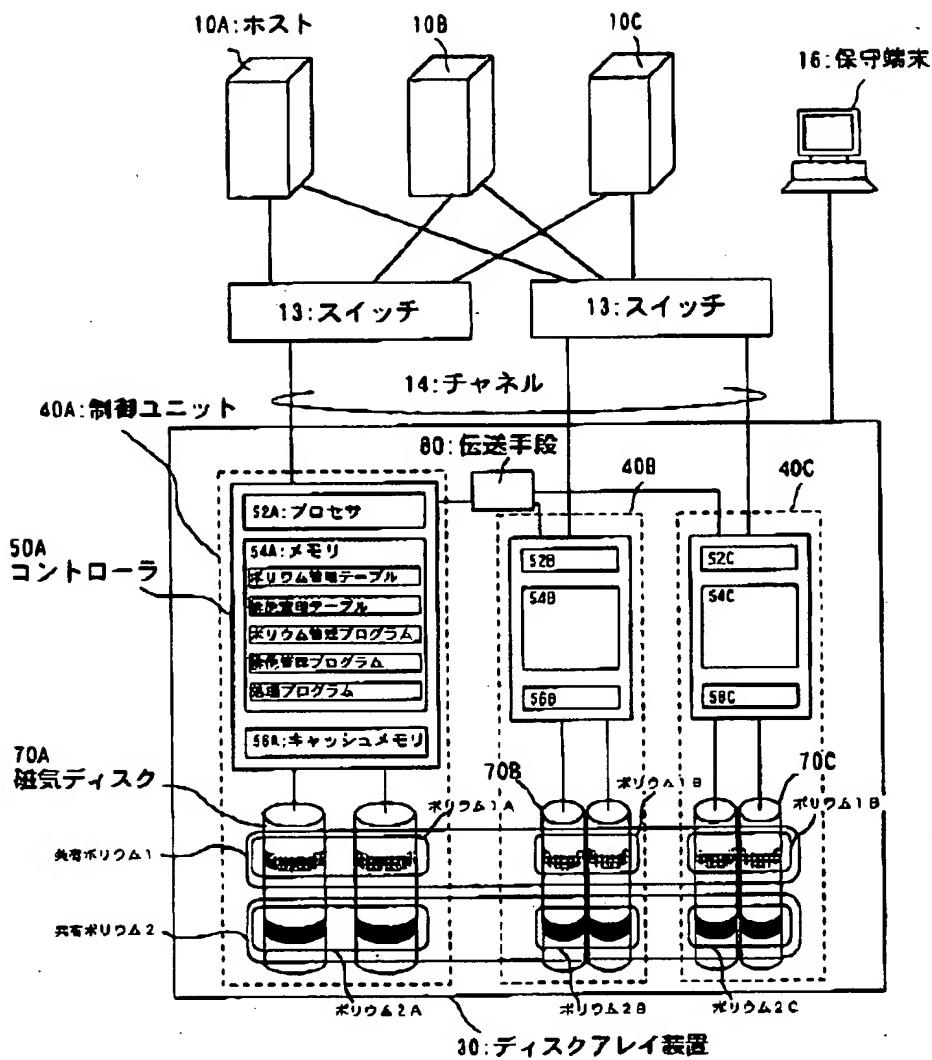
新規ボリューム チャネル番号 #B
ボリューム番号 2日

| 共有ボリューム 識別番号 | 対象ボリューム | |
|-----------------|---------|----------------|
| | チャネル番号 | ボリュームの 識別番号 |
| 共有ボリューム 2 | #A | ボリューム2A |
| | #B | ボリューム2B |
| | #C | ボリューム2C |

(18)

特開2003-241905

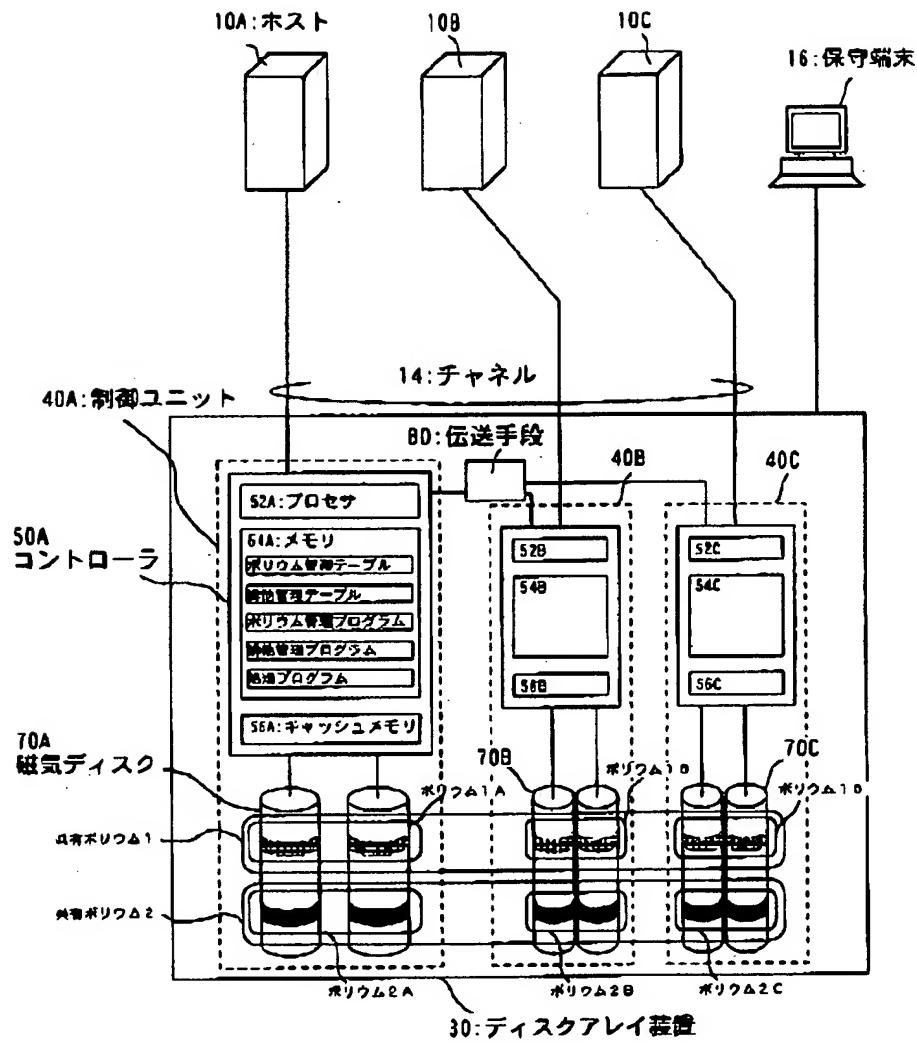
【図10】



(19)

特開2003-241905

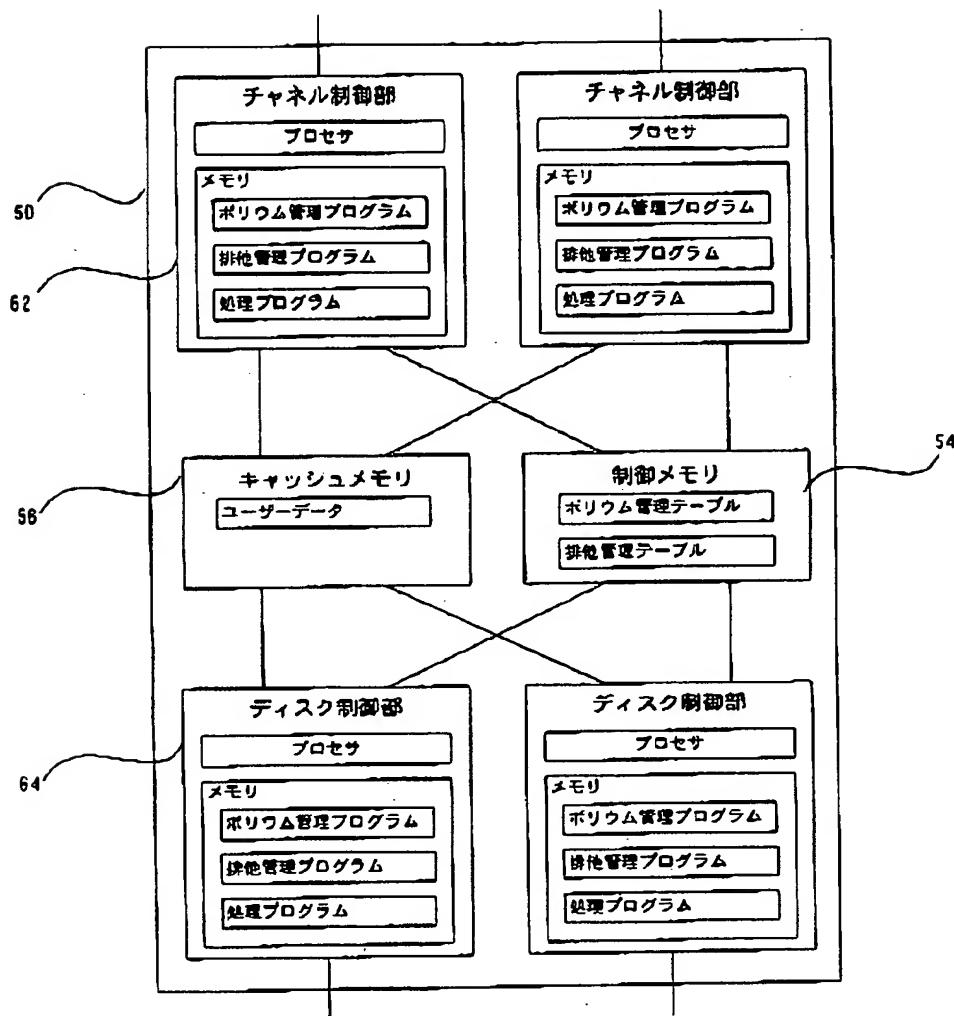
【図11】



(20)

特開2003-241905

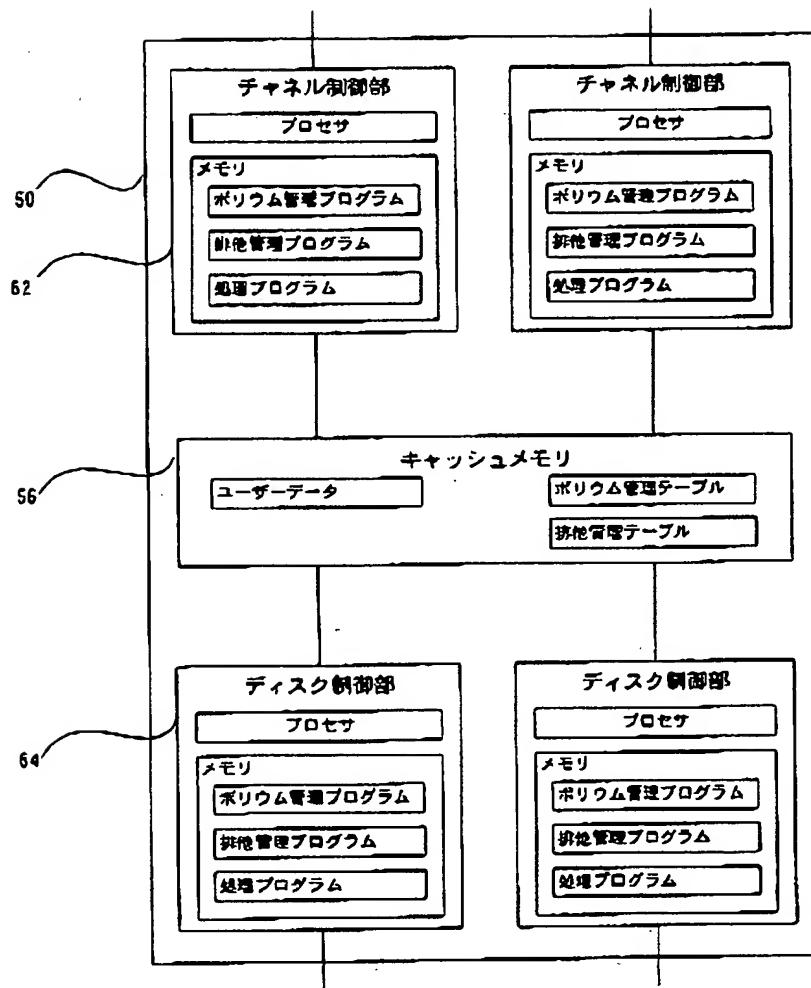
【図12】



(21)

特開2003-241905

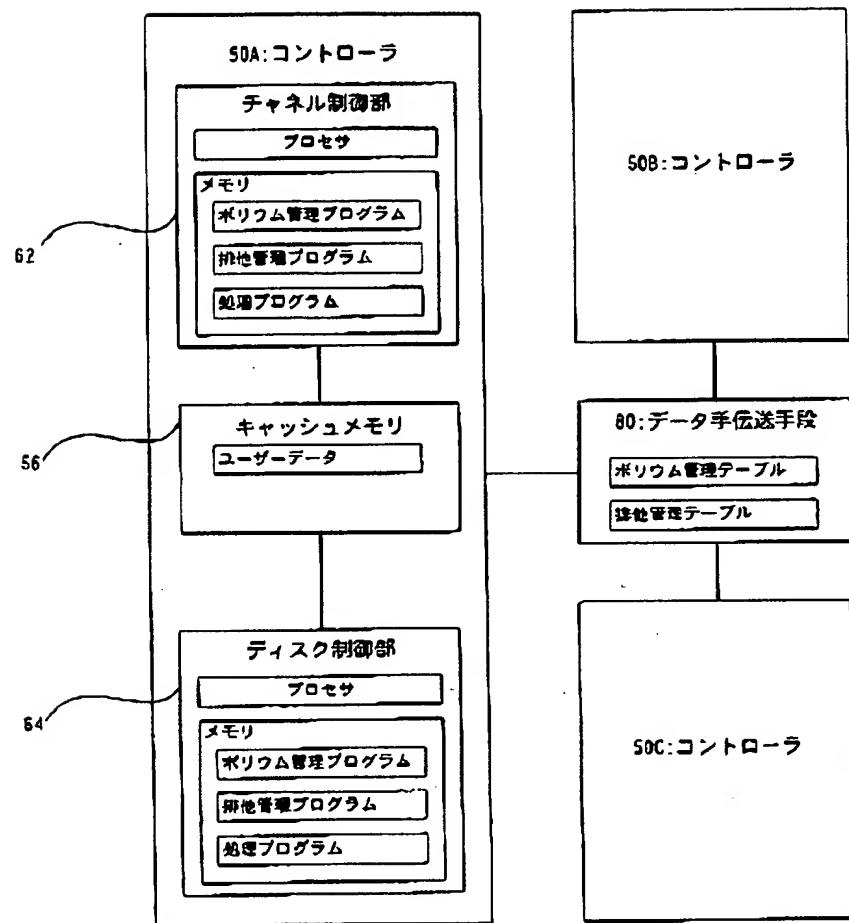
【図13】



(22)

特開2003-241905

【図14】



JP 2003-241905 A5 2005.8.18

【公報種別】特許法第17条の2の規定による補正の掲載

【部門区分】第6部門第3区分

【発行日】平成17年8月18日(2005.8.18)

【公開番号】特開2003-241905(P2003-241905A)

【公開日】平成15年8月29日(2003.8.29)

【出願番号】特願2002-39330(P2002-39330)

【国際特許分類第7版】

G 06 F 3/06

G 06 F 12/00

G 06 F 12/08

【F 1】

G 06 F 3/06 302Z

G 06 F 3/06 301B

G 06 F 12/00 545A

G 06 F 12/08 531C

G 06 F 12/08 557

【手続補正書】

【提出日】平成17年1月28日(2005.1.28)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

記憶手段と該記憶手段に対するデータの入出力を制御する制御装置とを有するユニットを複数備え、複数の該ユニットを接続するデータ伝送手段を備える記憶装置であって、
一のユニットの記憶手段に、他のユニットの記憶手段に記憶されたデータの複製を記憶し、

該一のユニットの制御装置が該他のユニットの記憶手段に記憶されたデータの読み出し要求を外部から受信したとき、該一のユニットの記憶手段に記憶した該複製を該外部に送信することを特徴とする記憶装置。

【請求項2】

請求項1に記載の記憶装置であって、
前記一のユニットの記憶手段に記憶されたデータの格納位置と、前記他のユニットの記憶手段に記憶されたデータの格納位置とを対応付けた情報を有し、

該一のユニットの制御装置が該他のユニットの記憶手段に記憶されたデータの読み出し要求を前記外部から受信したとき、該情報に基づいて、該読み出し要求を受けたデータに対応する格納位置に記憶された前記複製のデータを該外部に送信することを特徴とする記憶装置。

【請求項3】

請求項1又は2に記載の記憶装置であって、
前記他のユニットの制御手段は、該他のユニットの記憶手段に記憶された前記データに対する書き込み要求を前記外部から受信したとき、書き込み要求があった旨の通知を前記一のユニットに対して送信することを特徴とする記憶装置。

【請求項4】

請求項3に記載の記憶装置であって、
前記他のユニットの制御手段が、前記外部に対して、書き込み処理の終了を通知するのは

前記書き込み要求があった旨の通知を前記一の制御装置に送信した後であることを特徴とする制御装置。

【請求項 5】

請求項 1～4 のいずれかに記載の記憶装置であって、
前記他のユニットの記憶手段に記憶されたデータに対して書き込みを行うとき、
該他のユニットの制御装置は、前記一のユニットに対して、書き込みの内容を送信し、
該一のユニットの制御装置は、該他のユニットから、書き込みの内容を受信し、
該一のユニットの制御装置は、受信した書き込みの内容を前記複製に対して書き込むことを特徴とする記憶装置。

【請求項 6】

記憶手段と該記憶手段に対するデータの入出力を制御する制御装置とを有するユニットを複数備え、複数の該ユニットを接続するデータ伝送手段を備える記憶装置のデータ記憶方法であって、

一のユニットの記憶手段に、他のユニットの記憶手段に記憶されたデータの複製を記憶し、

該一のユニットの制御装置が該他のユニットの記憶手段に記憶されたデータの読み出し要求を外部から受信したとき、該一のユニットの制御装置が制御する記憶手段に記憶する該複製を該外部に送信することを特徴とするデータ記憶方法。

【請求項 7】

記憶手段に対するデータの入出力を制御し、他の記憶手段のデータの入出力を制御する他の制御装置とデータ伝送手段を介して接続される制御装置であって、

該他の記憶手段に記憶されたデータの複製を、該データ伝送手段を介して、自らが制御する該記憶手段に記憶し、

該他の記憶手段に記憶されたデータの読み出し要求を外部から受信したとき、該外部に対し、該複製を送信することを特徴とする制御装置。

【請求項 8】

請求項 7 に記載の制御装置であって、前記記憶手段を備えていることを特徴とする制御装置。

【請求項 9】

メモリと、磁気ディスクと、該磁気ディスクに対するデータの入出力を制御するコントローラとを有する複数の制御ユニットと、複数の該制御ユニットを接続するデータ伝送手段を備えるディスクアレイ装置であって、

一の制御ユニットの磁気ディスクのボリュームに、他の制御ユニットの磁気ディスクのボリュームに記憶されたデータの複製を記憶し、

該一の制御ユニットのメモリに、該複製を記憶している該他の制御ユニットの磁気ディスクのボリュームの識別番号と、該一の制御ユニットの磁気ディスクのボリュームの識別番号とを関連付けたデータであるボリューム管理テーブルを記憶し、

該一の制御ユニットのコントローラが、該他の制御ユニットの磁気ディスクに記憶されたデータの読み出し要求を外部から受信したとき、該一の制御ユニットのコントローラは、該ボリューム管理テーブルに基づいて、該読み出し要求に指定されたボリュームに記憶されているデータの複製が該一の制御ユニットの磁気ディスクに記憶されているかどうかを調べ、該複製が記憶されている場合には、該一の制御ユニットの磁気ディスクに記憶されている該複製のデータを該外部に送信し、該複製が記憶されていない場合には、該他の制御ユニットの磁気ディスクに記憶されているデータを該データ伝送手段を介して該一の制御ユニットのメモリに転送し、転送した該データを該外部に送信すること、

を特徴とするディスクアレイ装置。

【請求項 10】

メモリと、磁気ディスクと、該磁気ディスクに対するデータの入出力を制御するコントローラとを有する複数の制御ユニットと、複数の該制御ユニットを接続するデータ伝送手

(3)

JP 2003-241905 A5 2005.8.18

段を備えるディスクアレイ装置であって、

一の制御ユニットの磁気ディスクのボリュームに、他の制御ユニットの磁気ディスクのボリュームに記憶されたデータの複製を記憶し、

各制御ユニットのメモリに、排他処理の対象となる該磁気ディスクのボリュームを登録する排他管理テーブルを記憶し、

該一の制御ユニットのコントローラが、該他の制御ユニットの磁気ディスクのボリュームに対する書き込み要求を外部から受信したとき、該排他管理テーブルに、該書き込み要求に指定されたボリュームが登録されていない場合には、該他の制御ユニットに書き込み要求があった旨を通知し、各制御ユニットは、該排他管理テーブルに該書き込み先のボリュームを登録し、該一の制御ユニットのコントローラは、書き込みの内容を該データ伝送手段を介して該他の制御ユニットに送信し、各制御ユニットのコントローラは、該書き込みの内容を磁気ディスクに書き込み、該排他管理テーブルから該書き込み先のボリュームを削除することを特徴とするディスクアレイ装置。